(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 673 167 A2**

# (12) EUROPEAN PATENT APPLICATION

(21) Application number: **94118301.4**

(51) Int. Cl.[6]: **H04N 7/24**

(22) Date of filing: **21.11.94**

(30) Priority: **16.03.94 JP 46008/94**

(43) Date of publication of application:
**20.09.95 Bulletin 95/38**

(84) Designated Contracting States:
**DE FR GB NL**

(71) Applicant: **KABUSHIKI KAISHA TOSHIBA**
**72, Horikawa-cho,**
**Saiwai-ku**
**Kawasaki-shi,**
**Kanagawa-ken 210,**
**Tokyo (JP)**

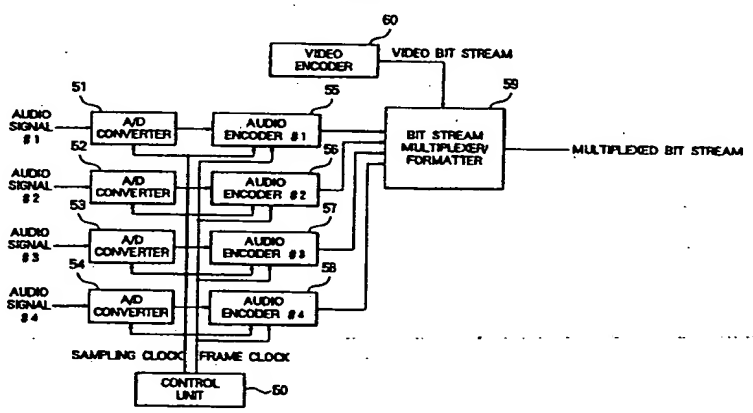
(72) Inventor: **Nakamura, Shin-ichi, c/o Intellectual Prop. Div.**
**Kabushiki Kaisha Toshiba,**
**1-1, Shibaura 1-chome**
**Minato-ku,**
**Tokyo (JP)**

(74) Representative: **Henkel, Feiler, Hänzel & Partner**
**Möhlstrasse 37**
**D-81675 München (DE)**

## (54) Signal encoding and decoding system allowing adding of signals in a form of frequency sample sequence upon decoding.

(57) In a signal compressing and encoding apparatus, a plurality of A/D converters sample input audio signals synchronously to produce time sample strings. An audio encoding circuit transforms time sample sequences in time/frequency by a frame interval to produce frequency sample sequences. The frequency sample sequences are compressed, encoded and formatted into a frame structure to generate audio bit streams. The audio bit streams are multiplexed by a formatter into a multiplexed bit stream. On the other hand, in a compressed signal decoding circuit, a quantization decoder decodes the frequency sample sequences for each bit stream separated and extracted selectively out of the multiplexed bit stream. A adding result the decoded frequency sample sequences to produce a summing result. The summing result is then transformed in frequency/time by a frequency/time mapping to produce the time sample sequences.

FIG. 1




EP 0 673 167 A2

The present invention relates to a signal compression and encoding apparatus for compressing and multiplexing, for example, audio or video signals and a compressed signal decoding apparatus for decoding the compressed and multiplexed audio or video signals. The present invention also relates to a combined system of these apparatuses.

To transmit/record digitized audio, acoustic signals requires a large volume of data, so that a huge transmission or storage capacity is required. With this respect, there has been a strong demand on a technique for compressing data.

Typically the human perceptual system is known not to receive all stimulation equally, and a part of the stimulation at that time is received while the remaining parts are not received. Discrimination of the perceived component and non-perceived component is known to be made effectively in frequency areas. With this respect, a compression algorithm of high efficiency using this principle has been developed. More specifically, time sample sequence is converted into a frequency sample sequence by using an orthogonal transform technique. The currently major method is the one to compress data on this frequency sample sequence. Examples of the transform in time/frequency include Discrete Cosine Transform (DCT) and Modified Discrete Cosine Transform (MDCT). The DCT algorithm is similar to Fast Fourier Transform (FFT), that converts data into sets of frequencies. In addition, subband filters where a large number of bandpass filters are aligned are also known.

One technique for compressing and encoding audio data is developed by the Moving Pictures Experts Group (MPEG), International Standards Organization (ISO) standard for compressing video (ISO/IEC-11172-3). Fig. 11 is a simple block diagram associated with this MPEG encoding. Input signals are audio digital signals sampled typically at sampling frequencies of 48 kHz, 44.1 kHz, or 32 kHz. Output can be selected for bit-streams of from 32 kbps to 448 kbps. Three layers, Layers I, II, and III, are defined according to compression efficiencies. Layers I and II based on the same fundamental algorithm are now described.

Input is supplied to a block 1 called mapping and then supplied to a subband filter group which segments a frequency region into thirty-two. The subband filter output is subjected to stripping to strip one sample by thirty-two samples. The subband filter is the one where it is ensured that the number of samples before passing through the subband filter is the same as that after passing therethrough without any change of the number of samples (i.e., the samples are critically sampled) and that the original time sample sequence can be de-formatted completely by this inverse transform. Accordingly, it can be understood that the input signal is transformed into a frequency sample sequence by the certain number of time samples in the approach using the subband filters, as in the one using the orthogonal transform technique.

As the mapping block 1, the input signal is supplied to a psychoacoustical model block 2. In this block 2, a frequency spectrum is obtained out of the input sample sequence by means of transformation in time/frequency (e.g., FFT) based on a certain frame independently of the mapping block 1. A length of this frame corresponds to 12 blocks (Layer I) or 36 blocks (Layer II) of the transform block of the above mentioned subband filter. In various schemes without using the subband filter, an orthogonal transform results in a large output. Accordingly, a frequency spectrum of this output may be used for calculation of the psychoacoustical model.

A signal to mask ratio (SMR) is then obtained between a signal level (S) of each subband, calculated out of the frequency spectrum by using the psychoacoustical model, and a mask level (M) which is not recognized acoustically due to masking effects. A quantization and coding block 3 resolves the output sample of the mapping into a product of the signal level (S) and a sample value (D(i)), namely,

$$\text{OUTPUT SAMPLE} = S * D(i).$$

The signal level (S) is referred to as a scale factor indicative of the level into which a sample of the maximum value is classified, within the subband of the frame. The frequency samples are quantized with a bit number obtained according to SMR into D'(i). The bit number is allocated to the frequency sample, and this process is referred to as a bit allocation. A quantization noise (N) or an artifact generated as a result of quantization will not be recognized even with the smaller bit number if this noise is lower than the masking level. The data is thus compressed.

The bit allocation, the scale factor, and the sample data are formatted into a bit stream sequence having a frame structure by a frame packing block 4. At the decoding side, this bit stream is received and frame synchronization is achieved by a frame unpack block 5. The bit allocation scale factor, and sample data are separated from each other for extraction. The frequency sample sequence is then reproduced by a quantization decoding block 6 and the time sample sequence is reconstituted by an inverse mapping block 7. As a result, a reproduced acoustic signal is obtained.

Fig. 12 shows an exemplified structure of a bit stream. A head of one frame is a header 21 having information about frame synchronization codes and other information regarding to, for example, mode. This is followed by the bit allocation 22, the scale factor 23, and the sample data 24. A supplementary data may follow it. One frame comprises one or more blocks for a time-to-frequency transform.

With a system for transmitting and storing audio signals associated with video images, the acoustic and audio signals should be in the form of two or more channels when two or more different languages are used. More specifically, the acoustic signal common for all languages and channels for respective languages are transmitted or stored. At the receiving or reproducing side, the common channel and one language are selected, which are summed and produced. Such "multi-channeled" acoustic signal may be compressed into bit streams by using one of following two techniques.

The first method is the one to compress all channels into a single bit stream. For MPEG, this approach has been realized in a new process of standardization as MPEG 2.

This method is, however, disadvantageous when the number of languages is increased. A bit rate of the bit stream (bit per second; bps) is increased to ensure sound of a significant quality. The bit rate (bps) of the compressed bit stream has certain limitations in a case where compression of two channels is fundamental as in MPEG. Such limitations are required to avoid elongated time for frame synchronization even with one audio frame has the increased number of bits in a two-channel mode. The limitation also ensures that the number of input buffer registers is not changed with the increased bit rate of the compressed bit stream during decoding operation. As apparent from the above, the standardized bit stream limits the number of channels available.

The second method is the one to construct a plurality of bit stream systems with signals on two or more channels being separate bit streams. This allows a system for many languages (channels) with the standard bit stream. In this event, only the bit stream on the desired channel is decoded at the decoding side. It is, however, necessary to decode separately the bit streams on the common acoustic channel and the channel for the selected language, increasing the amount of processing by two.

Fig. 13 shows an example of such multi-channel system. In this event, each audio input may be monaural or be on multiple channels as in stereo. In this system, a number of audio input signals are first subjected to audio compression/encoding by a number of audio encoders 31, 32, 33, 34. Bit streams are each multiplexed by a bit stream multiplexer/formatter 41 into a system bit stream. A video signal may be multiplexed, if the system include video signals. In addition, system information may also be multiplexed that is required for synchronization between the video and audio signals.

At the reproduction side, the video and audio bit streams and the system information are separated from each other by a demultiplexer 42. Processing of audio is made by means of selecting a desired bit stream from the reproduced audio bit streams and decoding it by using audio decoders 43 and 44. If a plurality of audio bit streams are selected, the audio decoders 43 and 44 decode compressed audio bit streams separately and independently. The decoded audio signals are mixed by a mixer/summer 45 and then produced.

As described above, conventionally, when a plurality of bit streams are formed for input signals of plural systems such as audio and acoustic signals to multiplex them, the multiplexed bit streams should each be decoded by using separate systems upon reproduction thereof. It is thus necessary to provide a plurality of decoding circuits having the same structure and function, disadvantageously increasing the scale and the dimension of the resultant circuit.

The present invention is to solve such problems and is directed to provide a signal compression and encoding apparatus, a compressed signal decoding apparatus, and a combined system of these apparatus that allow gathering frequency-to-time transform units together as one unit in a circuit structure of decoding system, thereby reducing a scale of the circuit when signals on plural input lines are each subjected to time-to-frequency transform, compressed, and multiplexed.

In the present invention, input signals on a plurality of input lines are each synchronized and sampled in sampling means to produce a time sample sequence. Next, each time sample sequence obtained by the sampling means is received in encoding means. The received time sample sequence is transformed in time/frequency while being synchronized by a predetermined time interval to produce frequency sample sequences. The frequency sample sequences are compressed and bit streams are produced that are formed of a predetermined format including the compressed frequency sample sequence. The above mentioned plurality of bit streams are multiplexed by multiplexing means.

A plurality of bit streams are selectively separated and extracted from the multiplexed bit stream by separating means. The frequency sample sequence is de-formatted by de-formatting means for each of these individual bit streams. The de-formatted frequency samples are summed to each other, the summing result of which is transformed in frequency/time to produce the time sample sequence.

More specifically, in the present invention, audio signals on plural input lines are sampled while being synchronized with each other to form a multiplexed bit stream that is frame-synchronized between audio on the individual input lines, which permits, upon decoding, summing in the frequency sample sequences as compared to summing of the time sample sequences conventionally. As a result, the frequency-to-time transform requires only one input line, allowing a simplified circuit structure.

These and other objects, features and advantages of the present invention will become more apparent in light of the following detailed description of the preferred embodiments as illustrated in the accompanying drawings in which:

Fig. 1 is a block diagram illustrating a structure of a signal compressing and encoding apparatus according to one embodiment of the present invention;

Fig. 2 is a block diagram illustrating a structure of a compressed signal decoding apparatus according to one embodiment of the present invention;

Fig. 3A is a view illustrating a relationship between a frequency spectrum and a time sample sequence in a frequency sample sequence of one bit stream;

Fig. 3B is a view illustrating a relationship between a frequency spectrum and a time sample sequence in a frequency sample sequence of one bit stream;

Fig. 3C is a view illustrating a relationship between a frequency spectrum and a time sample sequence in a frequency sample sequence of a added bit stream;

Fig. 4 shows an example where the transform block is overlapped with an adjacent block by 50%;

Fig. 5 is a circuit block diagram of a reproduction system for use in describing another embodiment of the present invention;

Fig. 6 is a view for use in describing and illustrating an asynchronous relation between a video frame and an audio frame;

Fig. 7 is a block diagram illustrating a structure of a signal compression and encoding apparatus according to another embodiment of the present invention;

Fig. 8 is a block diagram illustrating a structure of a compressed signal decoding apparatus according to another embodiment of the present invention;

Fig. 9 is a block diagram illustrating another embodiment in which a buffer is displaced in the compressed signal decoding apparatus illustrated in Fig. 8;

Fig. 10 is a block diagram illustrating a structure of yet another embodiment of the compressed signal decoding apparatus according to the present invention;

Fig. 11 is a block diagram for use in describing an audio signal compressing and encoding system;

Fig. 12 is a view showing a bit stream; and

Fig. 13 is a block diagram illustrating a structure of a conventional multi-channel compression system.

Embodiments of the present invention are described now with reference to the drawing.

Fig. 1 is a block diagram illustrating a structure of a signal compressing and encoding apparatus in a multiplexed bit stream transmission/reproduction system according to one embodiment of the present invention.

In this figure, 51, 52, 53, and 54 are A/D converters that receive audio input signals #1, #2, #3, and #4, respectively, on plural input lines to convert them into digital data. The audio input signals #1, #2, #3, and #4 may be music, effects sound, multilingual voice sounds, stereo sound, or monaural sound in media such as movies. 55, 56, 57, and 58 are audio encoders that are connected with outputs of the A/D converters 51, 52, 53, and 54, respectively.

The A/D converters 51-54 and the audio encoders 55-58 are supplied with a common sampling clock from a control unit 50. In addition, the audio encoders 55-58 are supplied with a frame clock from the control unit 50. The A/D converters 51-54 and the audio encoders 55-58 each produces an audio bit stream including frequency sample sequences for the audio input signals #1-#4 and sends them to a bit stream multiplexer/formatter 59.

More specifically, each of the A/D converters 51-54 samples the input audio signal at a timing of the above mentioned common sampling clock, and supplies the corresponding audio encoding unit 55-58 as the time sample sequence. Each of the audio encoding units 55-58 transforms in time/frequency the time sample sequences supplied from the corresponding A/D converters 51-54 by the frame intervals according to the above mentioned common frame clock to obtain the frequency sample sequences. They are compressed and encoded, and are then formatted into frame structures, thereby the audio bit streams are produced.

The bit stream multiplexer/formatter 59 multiplexes the audio bit streams produced by the audio encoders 55-58 and video bit streams taken from a video encoder 60 to produce a multiplexed bit stream.

Next, described is a compressed signal decoding apparatus in the multiplexed bit stream transmission/reproduction system according to this embodiment of the present invention.

Fig. 2 is a block diagram showing a structure of this compressed signal decoding apparatus. In this figure, 61 is a bit stream demultiplexing circuit 61 that receives the multiplexed bit streams to demultiplex the same into bit streams for individual input lines. This bit stream demultiplexing circuit 61 separates and extracts two or more desired audio bit streams out of the multiplexed each audio bit stream. 62 and 63 are frame synchronization circuits where each audio bit stream separated and extracted by the bit stream demultiplexing circuit 61 is subjected to frame synchronization and is then supplied to processing sections in a subsequent stage. 64 and 65 are de-formatting circuits where each frame-synchronized audio bit stream is subjected to analysis on each bit and is separated into the bit allocation, the scale factor, and the frequency sample sequences. 66 and 67 are quantization decoding units 66 that decode the frequency sample sequences of the individual audio bit streams. 68 is a common mixing/adding circuit 68. Each decoded frequency sample sequence is added for frequency samples at corresponding locations by this mixing/adding circuit 68 on one input line. The frequency sample sequence added on one input line is then supplied to a frequency/time mapping unit 69 and is converted into a time sample sequence for output. Thereafter, the time sample sequence is converted into an analog signal through a D/A converter 70 as an audio output signal.

Figs. 3 A-C illustrate a relationship of a frequency spectrum and a time sample sequence. In this event, Figs. 3A and 3B illustrate process of transforming frequency sample sequences consisting of different bit streams into time sample sequences by a certain transformation block. On the other hand, Fig. 3C illustrates process of transforming a summing result of the frequency sample sequences in Figs. 3A and 3B into a time sample sequence by a certain transformation block.

As mentioned above, the sum of the two or more frequency sample sequences can be transformed in frequency/time to obtain the original time sample sequence. This means that the sum of the frequency sample sequences is equivalent to the time sample sequence. The reason of this is described now.

The orthogonal transform and the inverse transform are generally expressed as follows:

$$X(m) = \sum_{n=0}^{N-1} k_m \, x(n) \, C_m^n \quad \ldots\ldots (1),$$

$$x(n) = \sum_{n=0}^{N-1} k_m \, X(m) \, C_m^n \quad \ldots\ldots (2),$$

where

x(n) represents an n-th time sample,

X(m) represents an m-th frequency sample,

$C_m^n$ represents a core of the orthogonal transform, or an orthogonal transform coefficient,

$$m = 0, \ldots, N-1$$
$$k_m = 1/2^{1/2} \quad m = 0,$$
$$= 1 \quad m \neq 0.$$

The equation (2) indicates the frequency-to-time transform. Let the frequency sample sequences of the bit stream be X(m) and Y(m), an output sample z(n) can be given by:

$$z(n) = \sum_{n=0}^{N-1} k_m \{ (aX(m) + bY(m)) \} \, C_m^n$$

$$= \sum_{n=0}^{N-1} k_m \, X(m) \, C_m^n + \sum_{n=0}^{N-1} k_m \, X(m) \, C_m^n$$

$$= ax(n) + by(n)$$

where a and b are summing coefficients.

Accordingly, the sum of the frequency sample sequences is equivalent to the time sample sequence.

For DCT, these equations are described in, for example, K. R. Rao and P. Yip, "Image Encoding Technique – DCT and International Standardization of It", Ohm Corporation.

According to this embodiment, the audio signals on plural input lines are sampled in a synchronous manner and a multiplexed bit stream is constructed with the audio signals for individual input lines in frame synchronization relation. As a result, the audio signals can be added upon decoding and reproduction in the form of the frequency sample sequences, which otherwise was made in the form of the time sample sequence in conventional arts. This requires only one input line for the frequency-to-time transform, resulting in a simplified circuit structure. In addition, the number of processing steps can be reduced when the above mentioned processing is achieved by software using a digital signal processor or the like.

In Fig. 3, a transform block may be overlapped with an adjacent block along the time axis for a case, for example, of the MDCT and the subband filters. Fig. 4 shows an example where the transform block is overlapped with an adjacent block by 50%. In this event, a window coefficient is multiplied with each frequency sample sequence. These frequency sample sequences are added together for an overlapping time. This frequency-to-time transform provides the time samples as twice the number as the frequency samples. However, the number of the samples becomes one-half as a result of the overlap adding processing. Accordingly, the number of the output samples become equal to that of the frequency samples. The present invention can thus be applied equally to this case.

A principle of the MDCT is described in, for example, "Analysis/Synthesis Filter Bank Design Based on Time Domain Aliasing Cancellation" (IEEE Transaction on Acoustics, Speech, and Signal Processing: Vol. ASSP-34, No. 5, October 1986, pages 1153 to 1161).

While this embodiment has thus been described in conjunction with the case where the frame synchronization is achieved for each audio bit stream, the frame synchronization may be achieved for at least one audio bit stream as shown in Fig. 5.

Next, a yet another embodiment of the present invention is described. The video frames and the audio frames are naturally asynchronous with each other as shown in Fig. 6. With this respect, this embodiment constructs such that the audio frames on a plurality of input lines are each in a synchronous relation and that the audio and video frames are also in a synchronous relation.

Fig. 7 is a block diagram illustrating a structure of a signal compression and encoding apparatus of this embodiment. In this figure, like reference numerals indicate like components and parts as in Fig. 1 and a detailed description thereof will be omitted. The signal compression and encoding apparatus in this embodiment is similar to the one described in Fig. 1 except that a video/audio synchronization information generating unit 71 is added. The video/audio synchronization information generating unit 71 is generates synchronization information for use in synchronizing the audio and video bit streams upon reproduction at a timing of the audio frames supplied from a predetermined audio coding unit (such as 55) as well as a timing of the video frames supplied from the video encoding unit 72. The video/audio synchronization information generating unit 71 supplies the synchronization information to the bit stream multiplexer/formatter 59. The synchronization information is further multiplexed into the multiplexed bit stream. This video/audio synchronization information is associated with a frame number of the audio signal to be reproduced at a given frame number of the video signal and with a sample number indicative of a position in the audio frame in question corresponding to an edge of the video frame.

Conventionally, it is necessary for producing video/audio synchronization information for each of the audio bit streams to achieve synchronization of the audio signals with the video signal by using the video/audio synchronization information because the audio bit streams in one input line are not in a synchronous relation with those in other input line. On the other hand, in the present embodiment, the audio bit streams in the individual input lines are in the synchronous relation with each other. The video/audio synchronization information is required to be generated only for one audio signal out of plural ones. Accordingly, the circuit required for this can be simplified.

Next, a compressed signal decoding apparatus thereof is described. Fig. 8 is a block diagram illustrating a structure of the compressed signal decoding apparatus. Like reference numerals indicate similar components and parts as in Fig. 2, and a detailed description thereof will be omitted. The bit stream demultiplexing circuit 61 demultiplexes the received multiplexed bit stream to extract the video bit stream, the selected audio bit stream, and the video/audio synchronization information. The video bit stream is supplied to a video decoding unit 81 and is decoded into a video output. The audio bit streams in the individual input lines are supplied to buffers 82 and 83. The audio bit streams in the individual systems are delayed by the buffers 82 and 83 and are supplied to a processing circuit in the subsequent stage (the frame synchronization circuit 62, the de-formatting circuits 64 and 65).

On the other hand, the video/audio synchronization information is supplied to a video/audio synchronization control circuit 84. The video/audio synchronization control circuit 84 calculates a delay amount for each of the buffers 82 and 83 according to the video/audio synchronization information, the video frame timing signal supplied from the video decoding unit 81, and the audio frame timing signal supplied from the frame synchronization circuit 62. More specifically, the video/audio synchronization control circuit 84 controls the delay amounts for the buffers 82 and 83 such that the audio frames of the individual systems are synchronized with the video frames. In this event, the audio bit streams in each system are synchronized with each other, so that one delay amount can be used commonly for the buffers 82 and 83. Accordingly, only one video/audio synchronization control circuit 84 is required for a simple structure.

While the buffers 82 and 83 are provided at the input for the audio bit stream in this embodiment, only one buffer 91 may be provided at the output for the audio bit stream as shown in Fig. 9 for synchronizing video and audio frames.

Yet another embodiment of the present invention is described. Fig. 10 is a block diagram illustrating a structure of a compressed signal decoding apparatus according to this embodiment. Like reference numerals indicate similar components and parts as in Fig. 2, and a detailed description thereof will be omitted. As shown in the figure, the compressed signal decoding apparatus comprises a synchronization determination unit 101. The synchronization determination unit 101 compares the frame timings detected from the individual audio bit streams to verify that the synchronization is achieved between the audio frames. If the synchronization is not achieved, the synchronization determination unit 101 carries out control to correct this synchronization error. For example, the synchronization determination unit 101 determines a correct synchronization timing by majority operation when the odd number of bit streams are selected. In addition, the synchronization determination unit 101 adjust the synchronization timing of minorities to the timing of the majorities. As a result, the synchronization of the audio bit streams are achieved with a higher accuracy.

The present invention can be applied to a case where the sampling frequencies are not equal in production of the bit streams of the audio signals in a plural systems. For example, the one sampling frequency may be one-half of the other as long as the two are in synchronism with each other. In this event, it is necessary to determine or select adequately a length of a time-to-frequency transform relative to the individual audio signals, and to make the position on the frequency axis of the frequency sample sequences of them be associated with each other in an adequate manner. In addition, the sampling frequencies may be in a ratio of, for example, 48 (kHz) to 32 (kHz) as long as they both are in synchronism with each other. In this event, the corresponding frequency samples can be obtained on the frequency axis as shown in Fig. 3 by means of calculating interpolating samples using a known frequency transform technique. The resultant frequency samples can be added on the frequency axis.

In the present invention, the number of the audio bit streams selected in the decoding apparatus not limited to two, and any positive integer can be applied. The number of the audio systems in each bit stream can also be selected arbitrary. The present invention is applicable as long as the mixing down can be achieved between the bit streams upon decoding.

As mentioned above, according to the present invention, the audio signals can be added upon decoding and reproduction in the form of the frequency sample sequences, which otherwise was made in the form of the time sample sequence in conventional arts. This requires only one system for the frequency-to-time transform, resulting in a simplified circuit structure.

In addition, when video and audio are multiplexed, the audio bit streams in the individual input lines are synchronized, so that the only thing required is to produce video/audio synchronization information for a single audio signal of a plurality of input lines. It is thus possible to simplify the circuit structure associated therewith.

## Claims

1. A system for encoding and decoding signals, the system transmitting or storing a multiplexed bit stream and decoding the transmitted or stored multiplexed bit stream, the multiplexed bit stream produced by compressing, encoding and multiplexing input signals of a plurality of input lines, and the system comprising:

   (a) a signal compressing and encoding apparatus comprising:

   sampling means (51-54) for sampling said input signals of said input lines while synchronizing between said input signals of said input lines, and outputting time sampled sequences composed of said sampled signals;

   encoding means (55-58) for transforming the time sampled sequences into frequency sampled

sequences while synchronizing between said time sampled sequences, compressing and encoding the transformed frequency sampled sequences, and producing a first bit stream having a predetermined format including the compressed and encoded frequency sampled sequences; and

multiplexing means (59) for multiplexing the first bit streams generated by the encoding means to output the multiplexed bit stream, and

(b) a compressed signal decoding apparatus comprising:

separating means (61) for separating selected first bit streams from the multiplexed bit streams, said first bit streams selected out of the multiplexed bit streams;

decoding means (66, 67) for decoding each of the first bit streams separated by the Separating means into the frequency sampled sequences;

adding means (68) for addition of the frequency sampled sequences decoded by the decoding means to produce a sum; and

transforming means (69) for transforming the sum into time sampled sequences, and outputting the transformed time sampled sequences.

2. A signal compressing and encoding apparatus for producing a multiplexed bit stream by compressing and encoding input signals of a plurality of input lines, comprising:

sampling means (51-54) for sampling said input signals of said input lines while synchronizing between said input signals of said input lines, and outputting time sampled sequences composed of said sampled signals;

encoding means (55-58) for transforming the time sampled sequences into frequency sampled sequences while synchronizing between said time sampled sequences, compressing and encoding the transformed frequency sampled sequences, and producing a first bit stream having a predetermined format including the compressed and encoded frequency sampled sequences; and

multiplexing means (59) for multiplexing the first bit streams generated by the encoding means to output the multiplexed bit stream.

3. A signal compressing and encoding apparatus, wherein sampling of said input signals of said input lines while synchronizing between said input signals of said input lines, and the time sampled Sequences are transformed into frequency sampled sequences while synchronizing between said time sampled sequences, and the transformed frequency sampled sequences are compressed and encoded, and a first bit stream having a predetermined format including the compressed and encoded frequency sampled sequences, and the produced bit stream is multiplexed, and the multiplexed bit stream is decoded,

further comprising:

separating means (61) for separating selected first bit streams from the multiplexed bit streams, said first bit streams selected out of the multiplexed bit streams;

decoding means (66, 67) for decoding each of the bit streams separated by the separating means into the frequency sampled sequences,

adding means (68) for addition of the frequency sampled sequences decoded by the decoding means to produce a sum, and

transforming means (69) for transforming the sum into the time sampled sequences, and outputting the transformed time sampled sequences.

4. A system as claimed in Claim 1 or an apparatus as claimed in Claim 3, wherein the decoding means (66, 67) decode the first bit streams selected by the separating means (61) into frequency sampled sequences while synchronizing with only one first bit stream of the separated first bit streams.

5. A System for encoding and decoding signals as claimed in Claim 1, wherein the signal compressing and encoding apparatus further comprises synchronization information generating means (71) for generating synchronization information for use in achieving synchronization with the first bit stream generated by the encoding means (55-58) with a second bit stream which is different from the first bit stream; the multiplexing means (59) for multiplexing the first bit stream, the second bit stream, and the synchronization information to output a multiplexed bit stream; and the Separating means (61) for separating selected first bit streams, second bit streams and synchronization information selected from the multiplexed bit streams,

and the compressed signal decoding apparatus further comprises synchronization controlling means (84) for ensuring synchronization between the first and the second bit streams in accordance

with the separated synchronization information.

6. A signal compressing and encoding apparatus as claimed in Claim 2, further comprising:
synchronization information generating means (71) for generating synchronization information for use in achieving synchronization between the first bit stream generated by the encoding means (55-58) and a second bit stream which is different from the first bit stream, wherein multiplexing means (59) multiplexes the first bit stream, the second bit stream and the synchronization information to output multiplexed streams.

7. A signal compressing and decoding apparatus as claimed in Claim 3, wherein the Separating means (61) separates selected first bit streams, second bit streams which are different from the first bit streams, and the synchronization information, and further comprising synchronization controlling means (84) for ensuring synchronization between the first and the second bit streams in accordance with the separated synchronization information.

8. A system as claimed in Claim 5 or an apparatus as claimed in Claim 7, wherein the synchronization controlling means (84) ensures synchronization between the first and the second bit streams at a position between the separating means (61) and the decoding means (66, 67).

9. A system as claimed in Claim 1 or an apparatus as claimed in Claim 3, wherein the synchronization controlling means (84) ensures synchronization between the first and the second bit streams after transformation carried out by the transforming means (69).

10. A system as claimed in Claim 1 or an apparatus as claimed in Claim 3, further comprising:
synchronization detecting means (62, 63) for detecting a frame synchronization timing for the first bit stream separated by the separating means (61); and
controlling means (101) for comparing the detected frame synchronization timings with each other to determine a synchronization error between the first bit streams, the controlling means carrying out control such that the synchronization error is eliminated when the synchronization error is caused.

11. A System or an apparatus as claimed in Claims 5 - 9, wherein the first bit stream is a bit stream relating to audio and the second bit stream is a bit stream relating to video.

12. A method for transmitting and storing a multiplexed bit stream and decoding the transmitted or stored multiplexed bit stream, the multiplexed bit stream produced by comprising, encoding and multiplexing input signals of a plurality of input lines, and the method comprising the steps of:
sampling said input signals of said input lines while synchronizing between said input signals of said input lines, and outputting time sampled sequences composed of said sampled signals;
transforming the time sampled sequences into frequency sampled sequences while synchronizing between said time sampled sequences, compressing and encoding the transformed frequency sampled sequences, and producing a first bit stream having a predetermined format including the compressed and encoded frequency sampled sequences;
multiplexing the first bit streams generated by the encoding means to output the multiplexed bit stream;
separating selected first bit streams from the multiplexed bit streams, said first bit streams selected out of the multiplexed bit streams;
decoding each of the first bit streams separated by the separating means into the frequency sampled sequences;
adding means for addition of the frequency sampled sequences decoded by the decoding means to produce a sum;
transforming the sum into the time sampled sequence, and outputting the transformed time sampled sequence.

FIG. 1

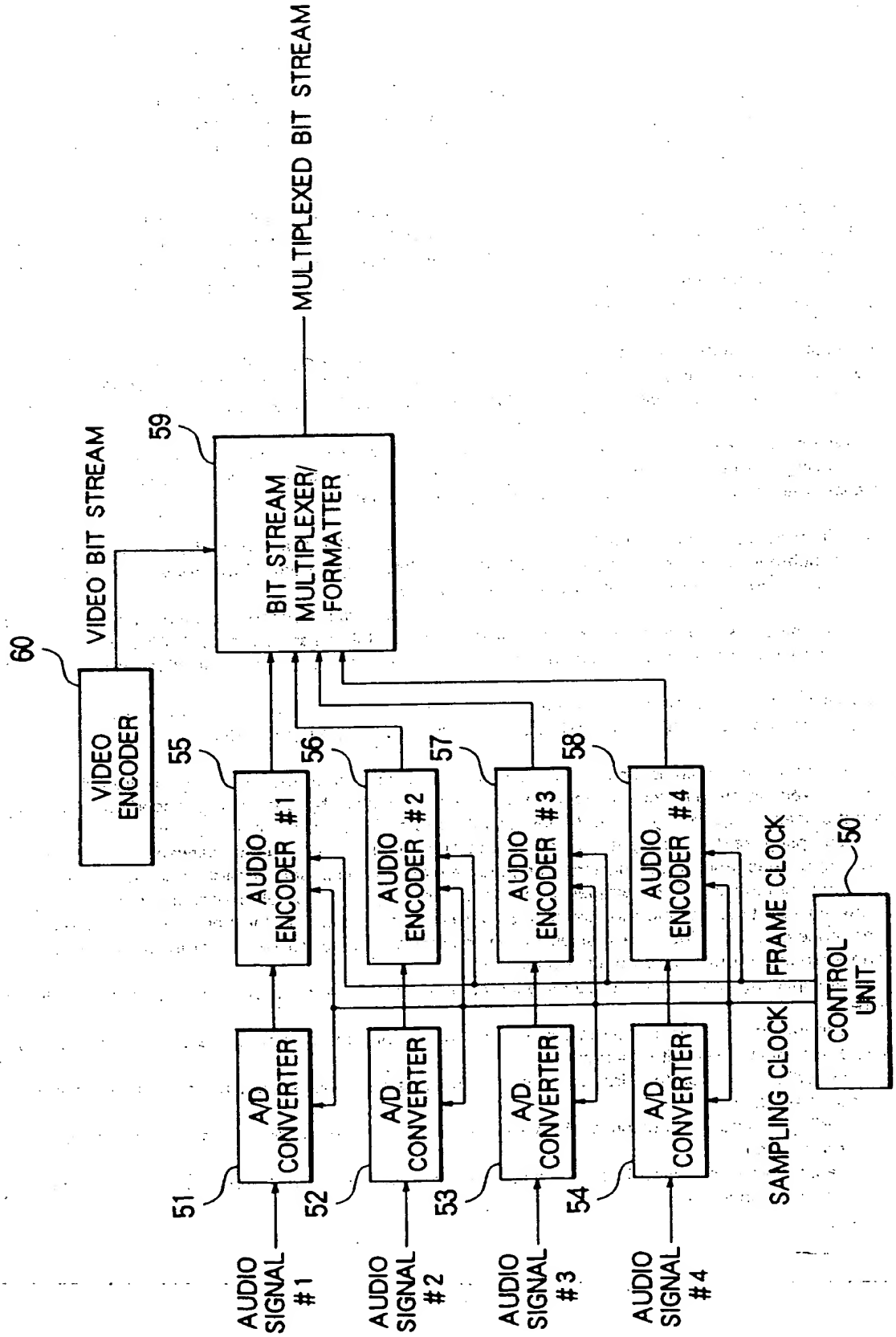

MULTIPLEXED BIT STREAM
59
BIT STREAM MULTIPLEXER/ FORMATTER
VIDEO BIT STREAM
60
VIDEO ENCODER
55
AUDIO ENCODER #1
56
AUDIO ENCODER #2
57
AUDIO ENCODER #3
58
AUDIO ENCODER #4
FRAME CLOCK
50
CONTROL UNIT
SAMPLING CLOCK
51
A/D CONVERTER
52
A/D CONVERTER
53
A/D CONVERTER
54
A/D CONVERTER
AUDIO SIGNAL #1
AUDIO SIGNAL #2
AUDIO SIGNAL #3
AUDIO SIGNAL #4

FIG. 2

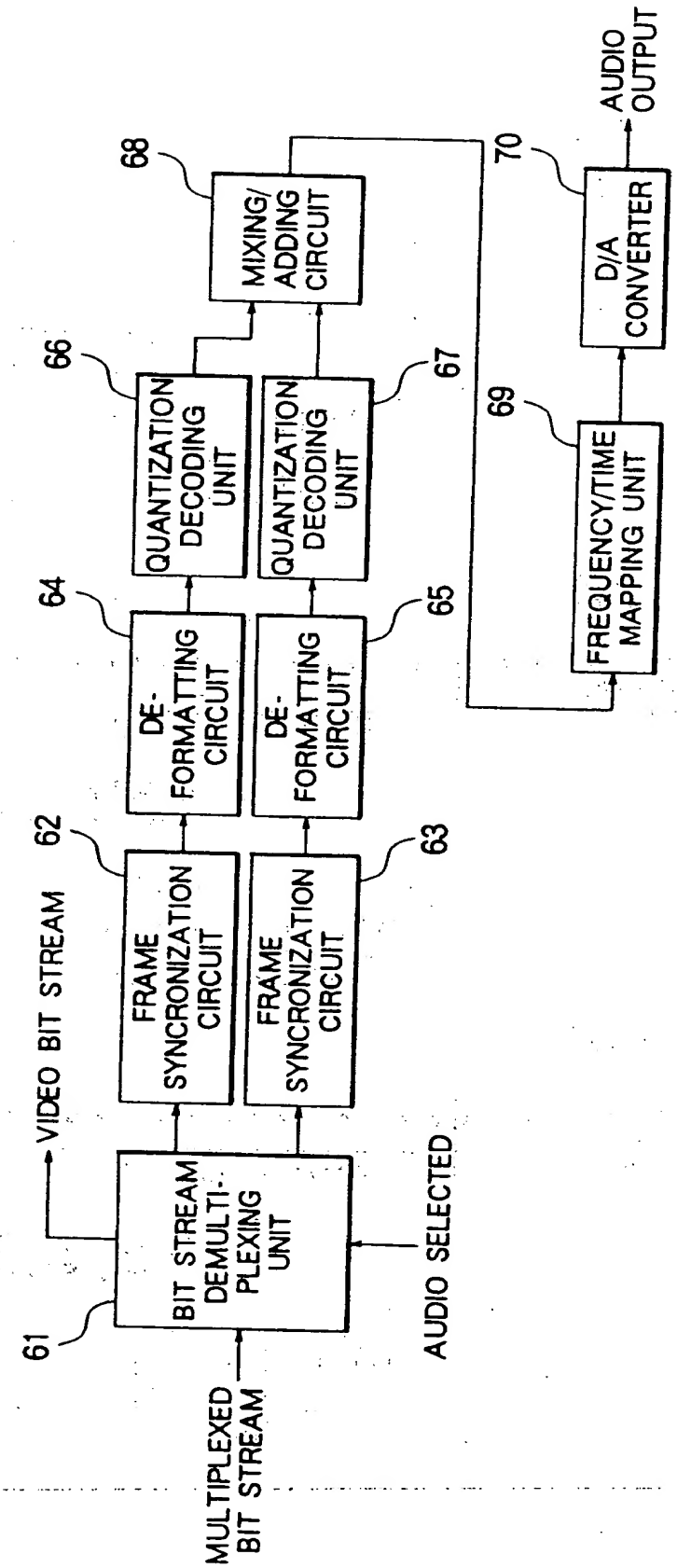

MULTIPLEXED BIT STREAM
61
BIT STREAM DEMULTI-PLEXING UNIT
VIDEO BIT STREAM
AUDIO SELECTED
62
FRAME SYNCRONIZATION CIRCUIT
63
FRAME SYNCRONIZATION CIRCUIT
64
DE-FORMATTING CIRCUIT
65
DE-FORMATTING CIRCUIT
66
QUANTIZATION DECODING UNIT
67
QUANTIZATION DECODING UNIT
68
MIXING/ ADDING CIRCUIT
69
FREQUENCY/TIME MAPPING UNIT
70
D/A CONVERTER
AUDIO OUTPUT

FIG. 3A

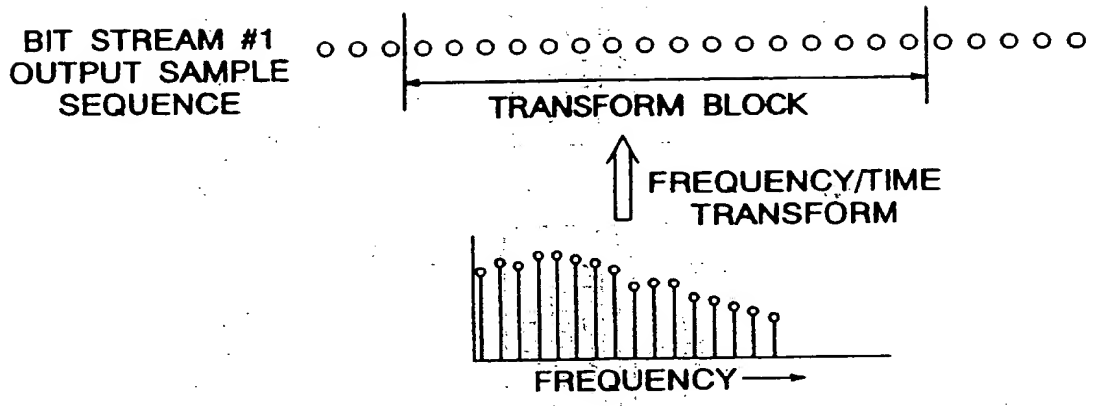



FIG. 3B

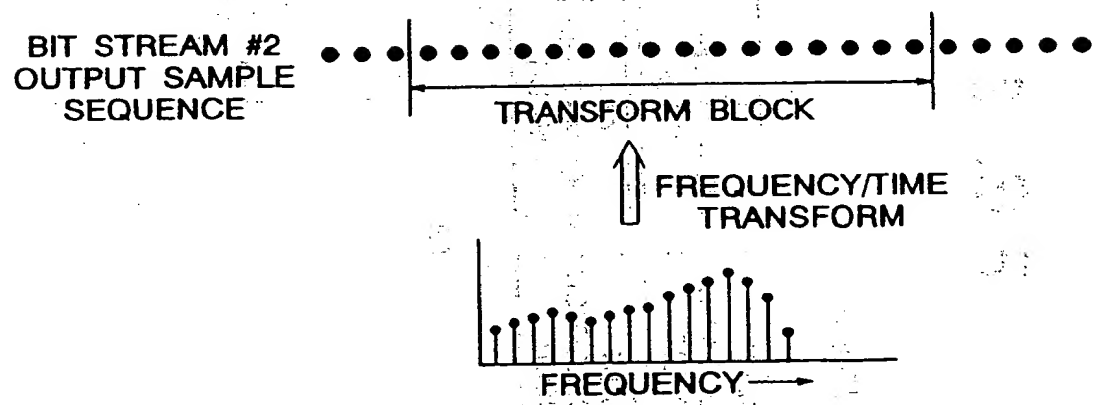



FIG. 3C

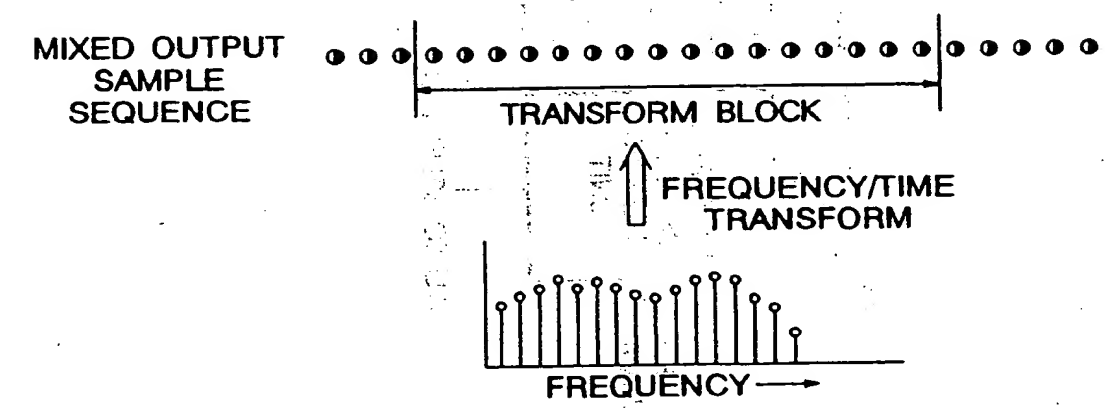


FIG. 4

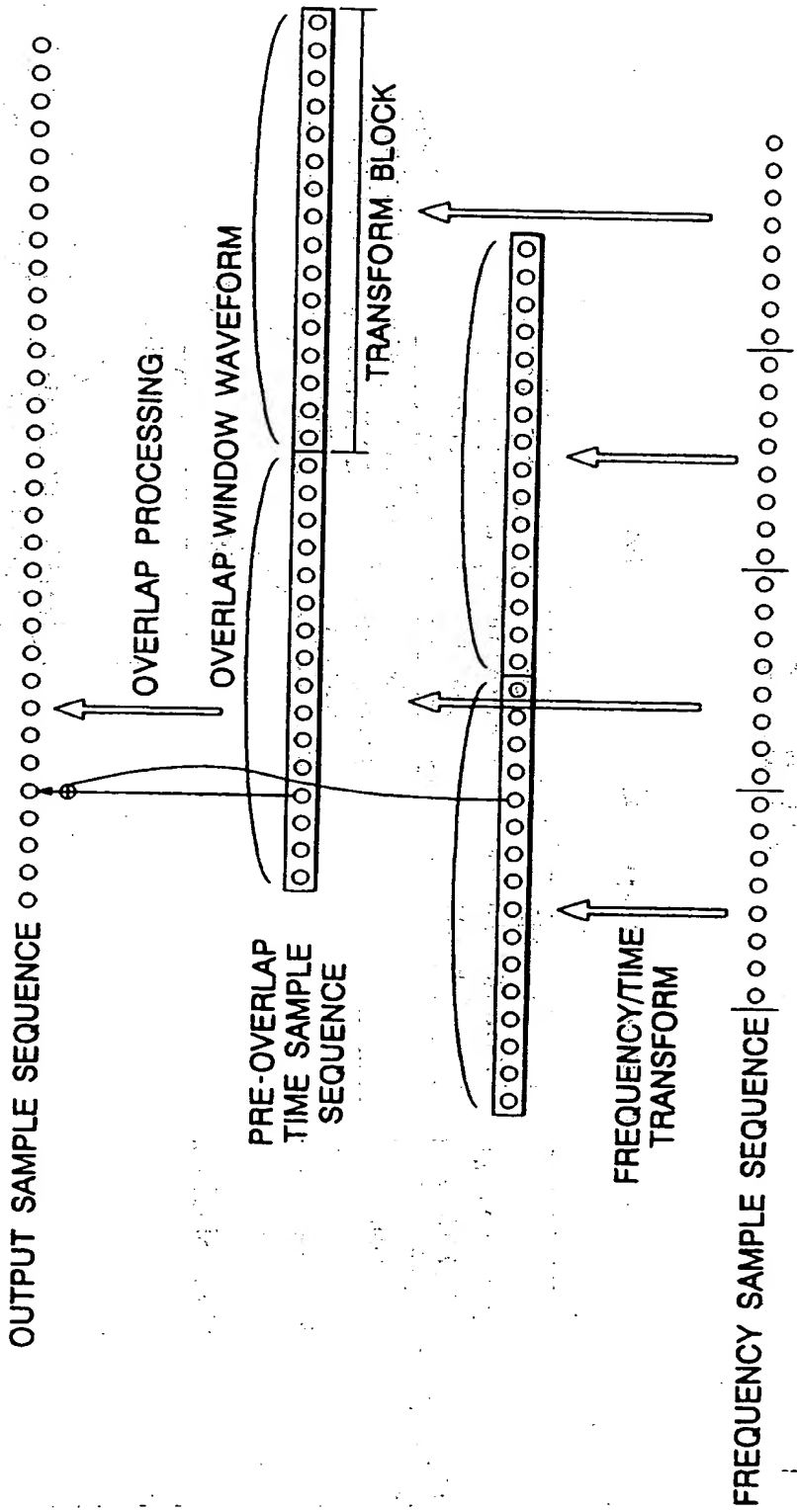

OUTPUT SAMPLE SEQUENCE
OVERLAP PROCESSING
OVERLAP WINDOW WAVEFORM
TRANSFORM BLOCK
PRE-OVERLAP TIME SAMPLE SEQUENCE
FREQUENCY/TIME TRANSFORM
FREQUENCY SAMPLE SEQUENCE

FIG. 5

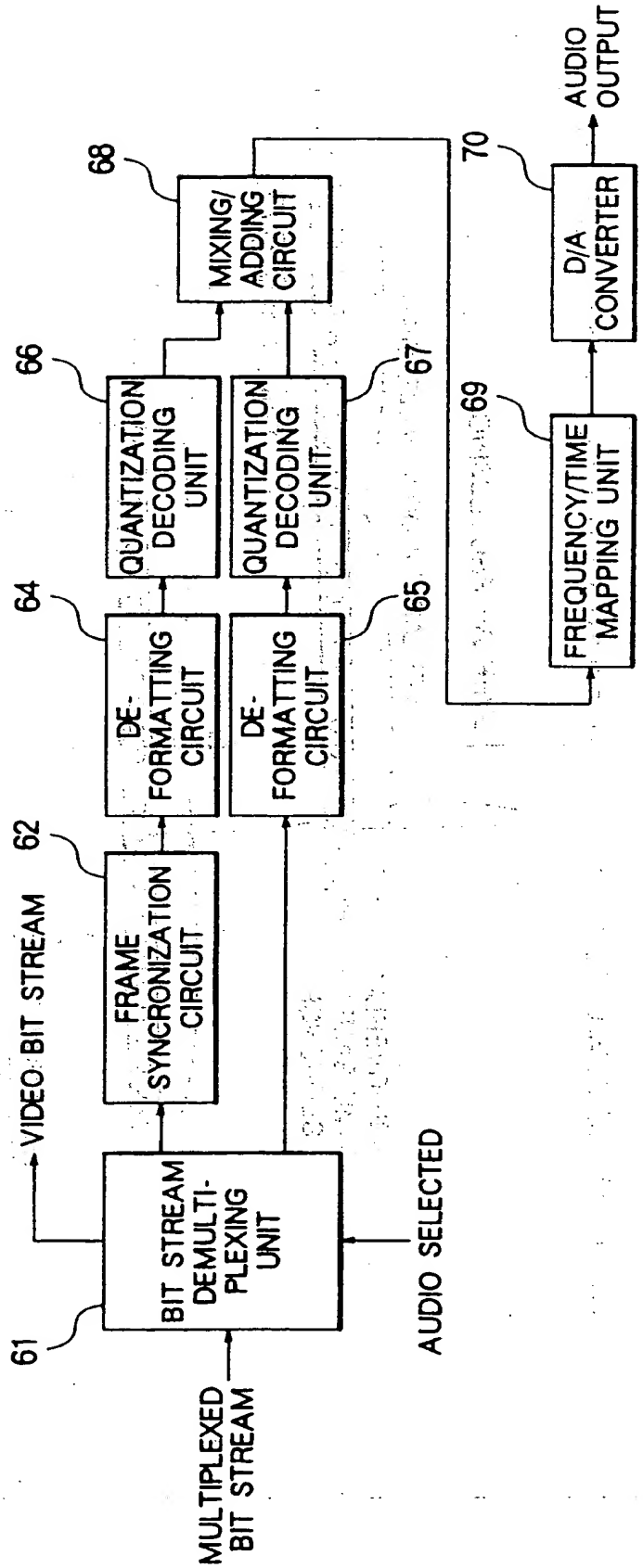

MULTIPLEXED BIT STREAM
61
BIT STREAM DEMULTI-PLEXING UNIT
VIDEO BIT STREAM
AUDIO SELECTED
62
FRAME SYNCRONIZATION CIRCUIT
64
DE-FORMATTING CIRCUIT
65
DE-FORMATTING CIRCUIT
66
QUANTIZATION DECODING UNIT
67
QUANTIZATION DECODING UNIT
68
MIXING/ ADDING CIRCUIT
69
FREQUENCY/TIME MAPPING UNIT
70
D/A CONVERTER
AUDIO OUTPUT

FIG. 6

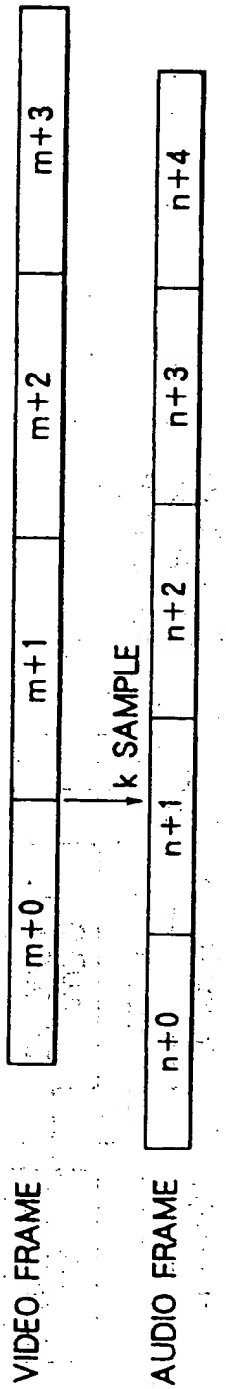


FIG. 7

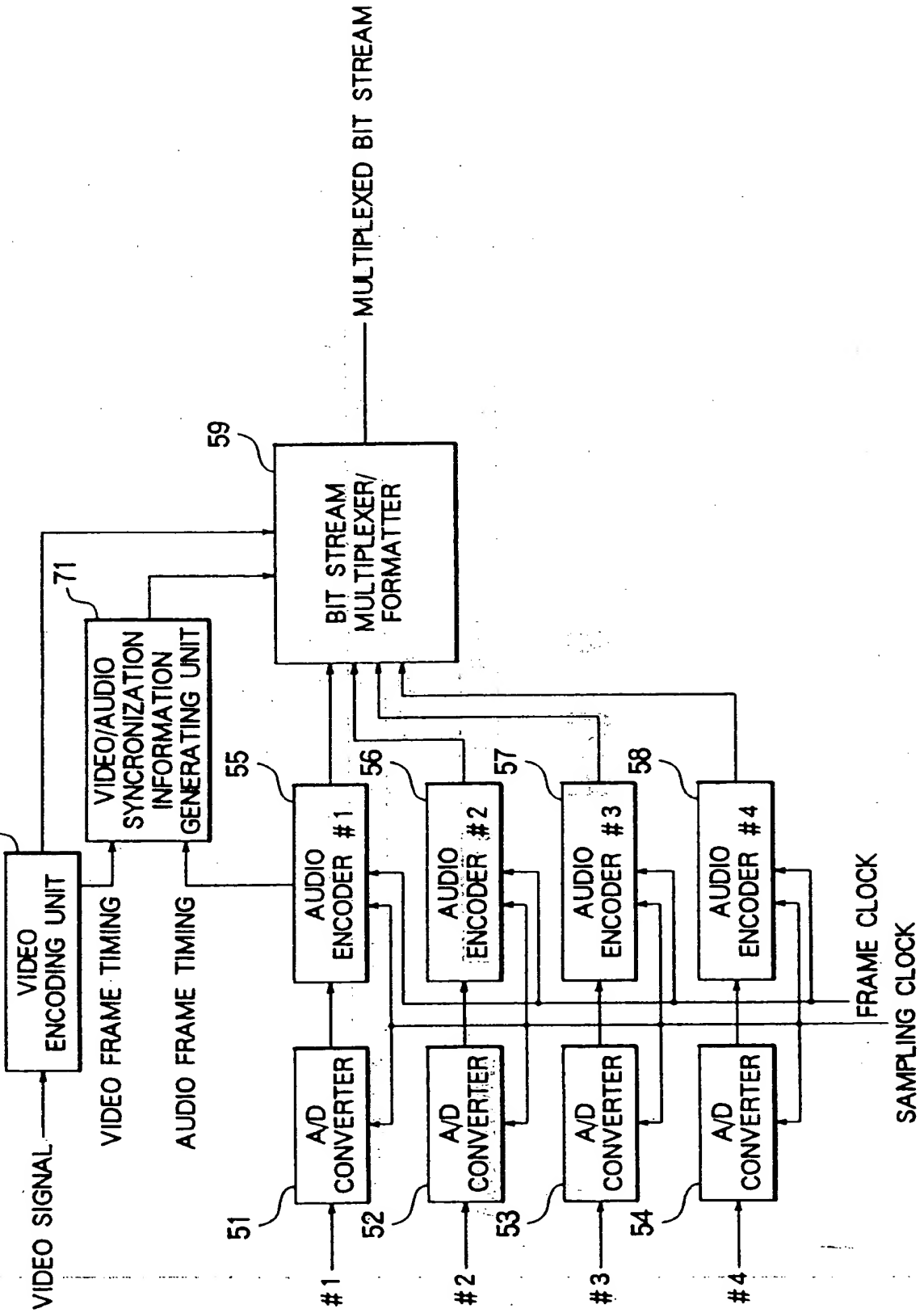

VIDEO SIGNAL
VIDEO ENCODING UNIT
72
VIDEO FRAME TIMING
VIDEO/AUDIO SYNCRONIZATION INFORMATION GENERATING UNIT
71
AUDIO FRAME TIMING
BIT STREAM MULTIPLEXER/ FORMATTER
59
MULTIPLEXED BIT STREAM
AUDIO ENCODER #1
AUDIO ENCODER #2
AUDIO ENCODER #3
AUDIO ENCODER #4
55
56
57
58
A/D CONVERTER
A/D CONVERTER
A/D CONVERTER
A/D CONVERTER
51
52
53
54
#1
#2
#3
#4
FRAME CLOCK
SAMPLING CLOCK

FIG. 8

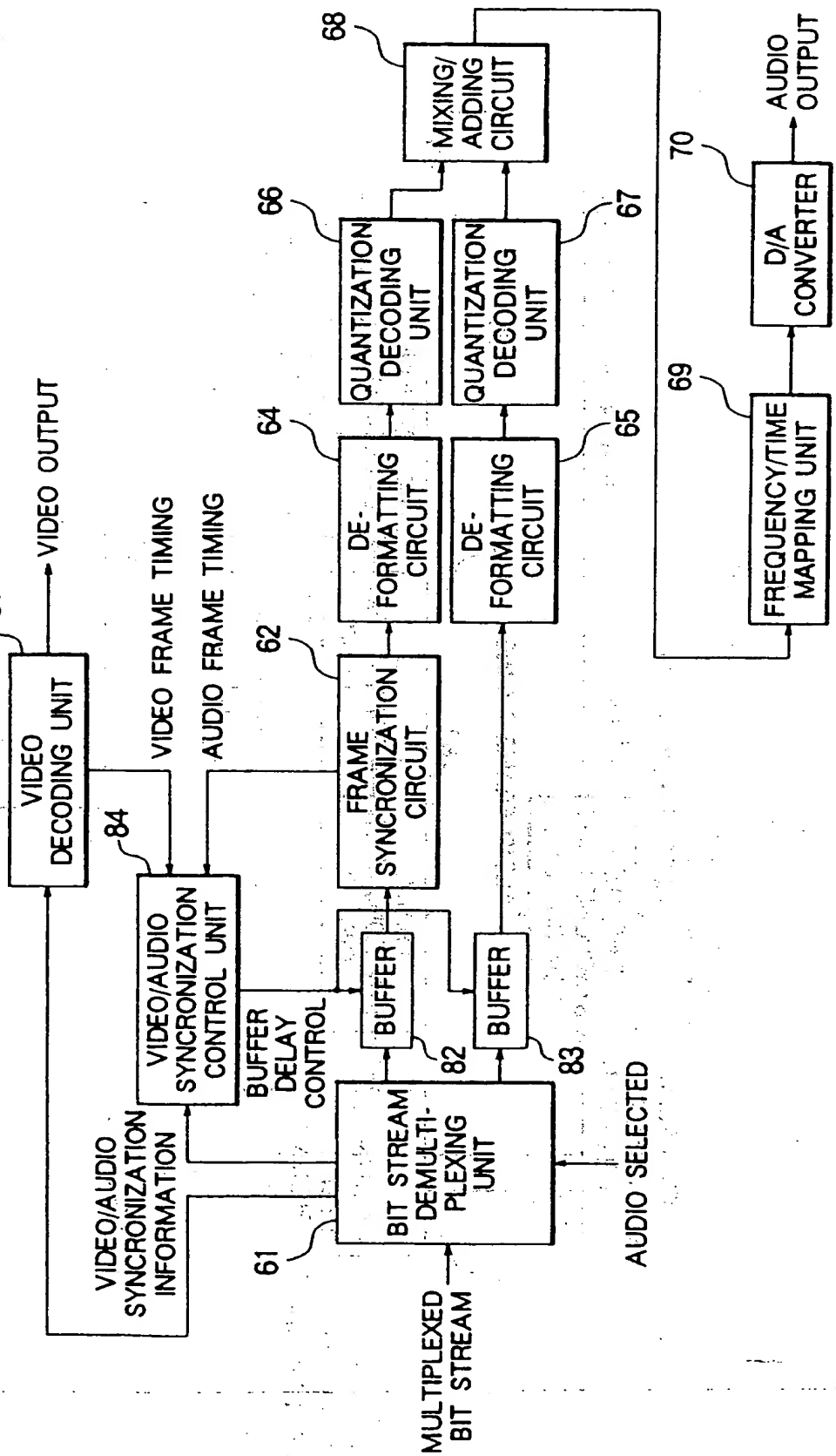

MULTIPLEXED BIT STREAM
BIT STREAM DEMULTI-PLEXING UNIT
61
VIDEO/AUDIO SYNCRONIZATION INFORMATION
BUFFER DELAY CONTROL
VIDEO/AUDIO SYNCRONIZATION CONTROL UNIT
84
BUFFER
82
BUFFER
83
FRAME SYNCRONIZATION CIRCUIT
62
DE-FORMATTING CIRCUIT
64
DE-FORMATTING CIRCUIT
65
QUANTIZATION DECODING UNIT
66
QUANTIZATION DECODING UNIT
67
MIXING/ ADDING CIRCUIT
68
AUDIO SELECTED
FREQUENCY/TIME MAPPING UNIT
69
D/A CONVERTER
70
AUDIO OUTPUT
VIDEO DECODING UNIT
81
VIDEO FRAME TIMING
AUDIO FRAME TIMING
VIDEO OUTPUT

FIG. 9

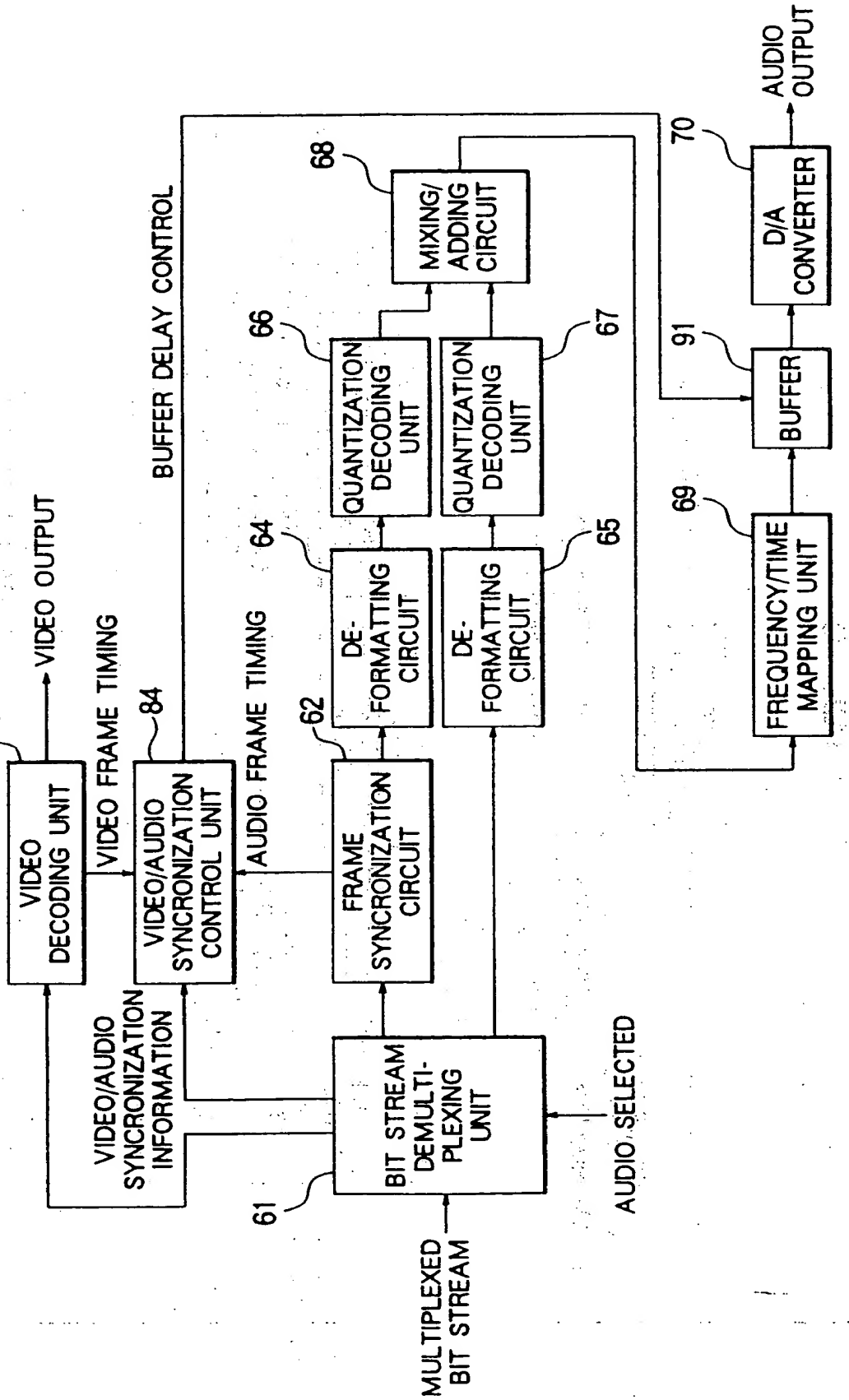

MULTIPLEXED BIT STREAM
61
BIT STREAM DEMULTI-PLEXING UNIT
AUDIO SELECTED
VIDEO/AUDIO SYNCRONIZATION INFORMATION
FRAME SYNCRONIZATION CIRCUIT
62
DE-FORMATTING CIRCUIT
64
DE-FORMATTING CIRCUIT
65
QUANTIZATION DECODING UNIT
66
QUANTIZATION DECODING UNIT
67
MIXING/ ADDING CIRCUIT
68
FREQUENCY/TIME MAPPING UNIT
69
BUFFER
91
D/A CONVERTER
70
AUDIO OUTPUT
VIDEO/AUDIO SYNCRONIZATION CONTROL UNIT
84
AUDIO FRAME TIMING
VIDEO FRAME TIMING
VIDEO DECODING UNIT
81
VIDEO OUTPUT
BUFFER DELAY CONTROL

FIG. 10

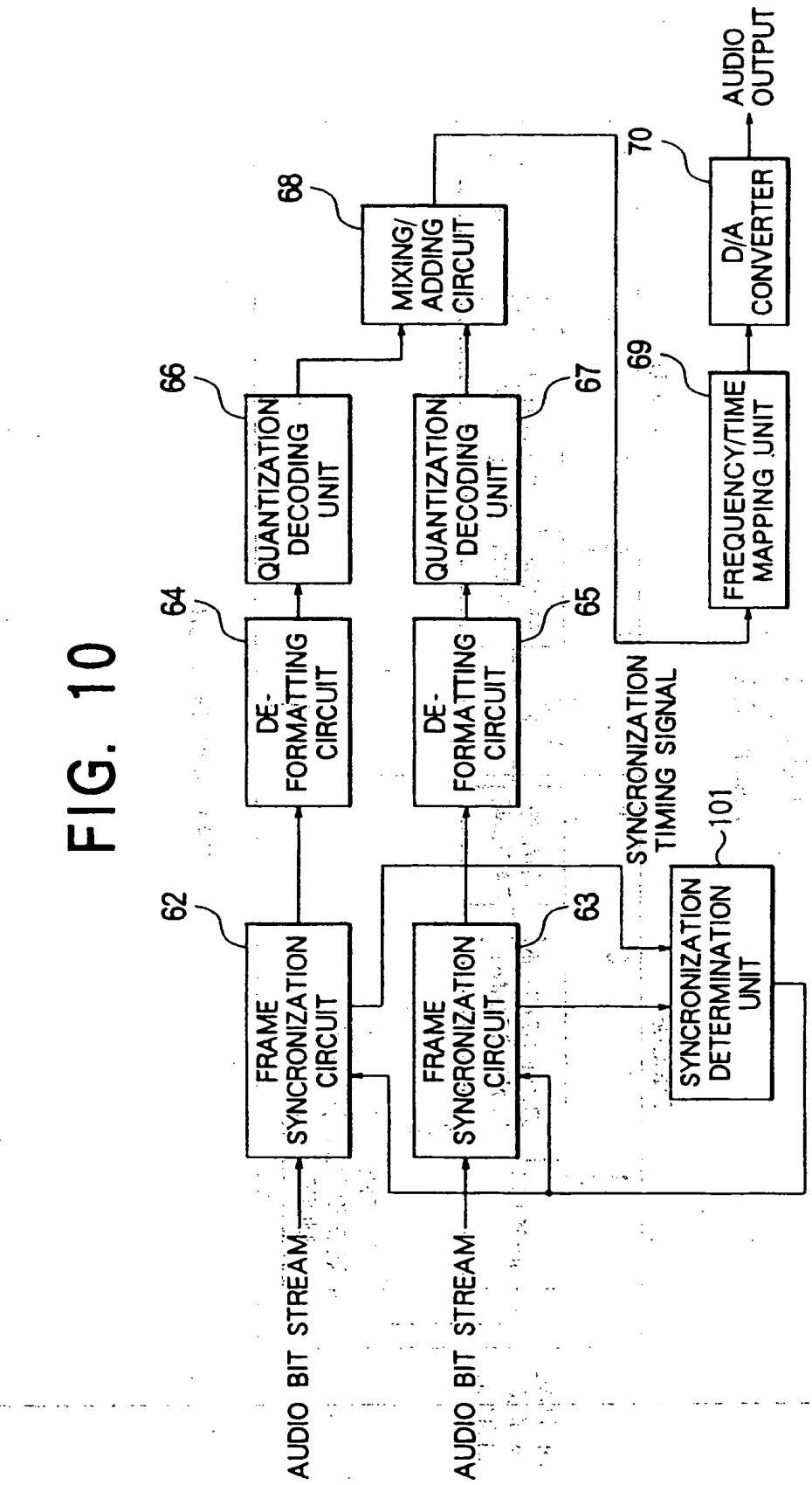

AUDIO BIT STREAM
AUDIO BIT STREAM
FRAME SYNCRONIZATION CIRCUIT
62
FRAME SYNCRONIZATION CIRCUIT
63
DE-FORMATTING CIRCUIT
64
DE-FORMATTING CIRCUIT
65
QUANTIZATION DECODING UNIT
66
QUANTIZATION DECODING UNIT
67
MIXING/ ADDING CIRCUIT
68
FREQUENCY/TIME MAPPING UNIT
69
D/A CONVERTER
70
AUDIO OUTPUT
SYNCRONIZATION TIMING SIGNAL
SYNCRONIZATION DETERMINATION UNIT
101

FIG. 11
PRIOR ART

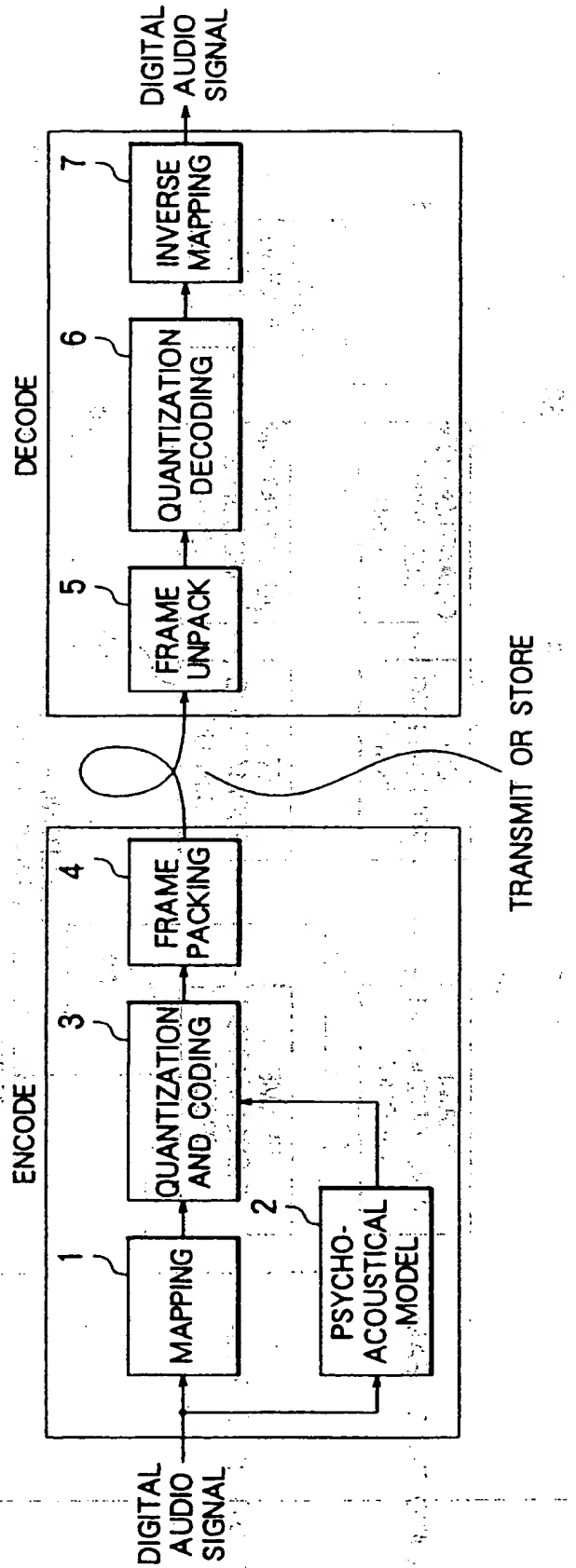


FIG. 12
PRIOR ART

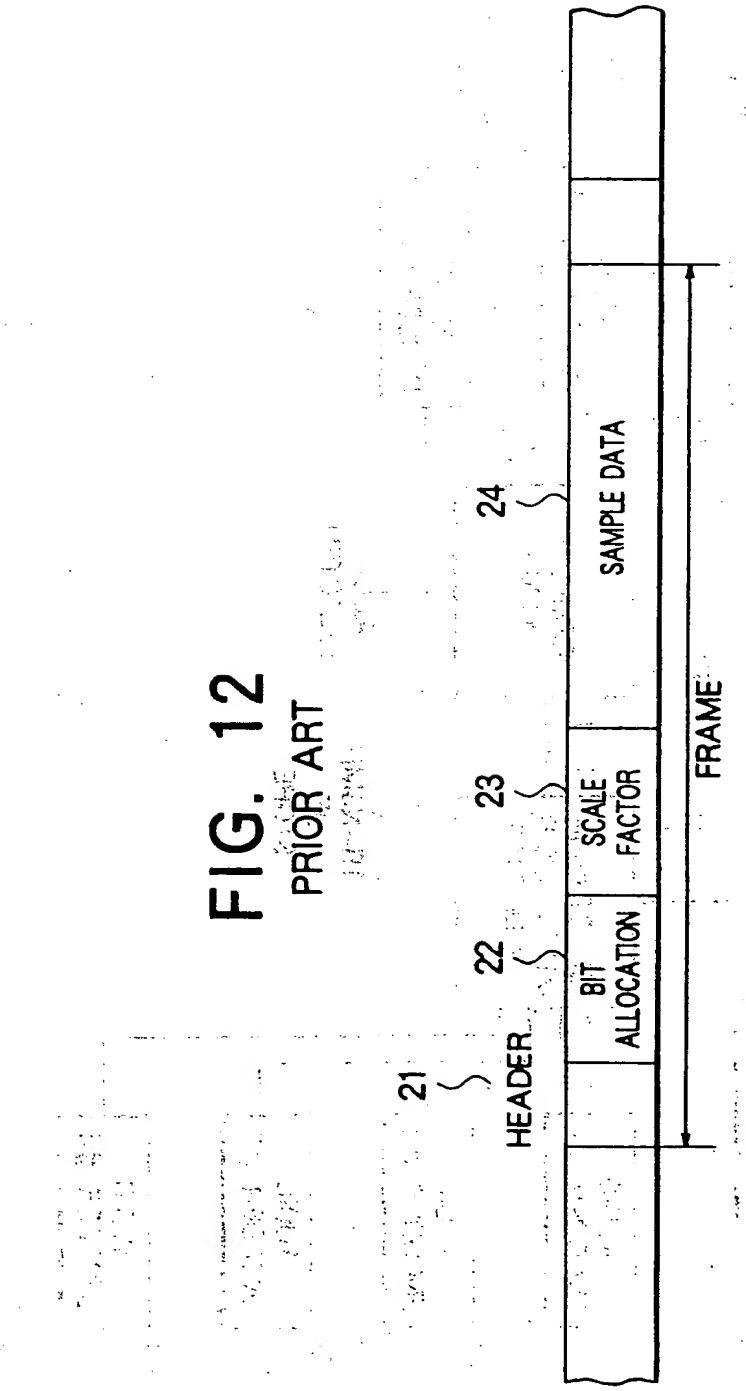
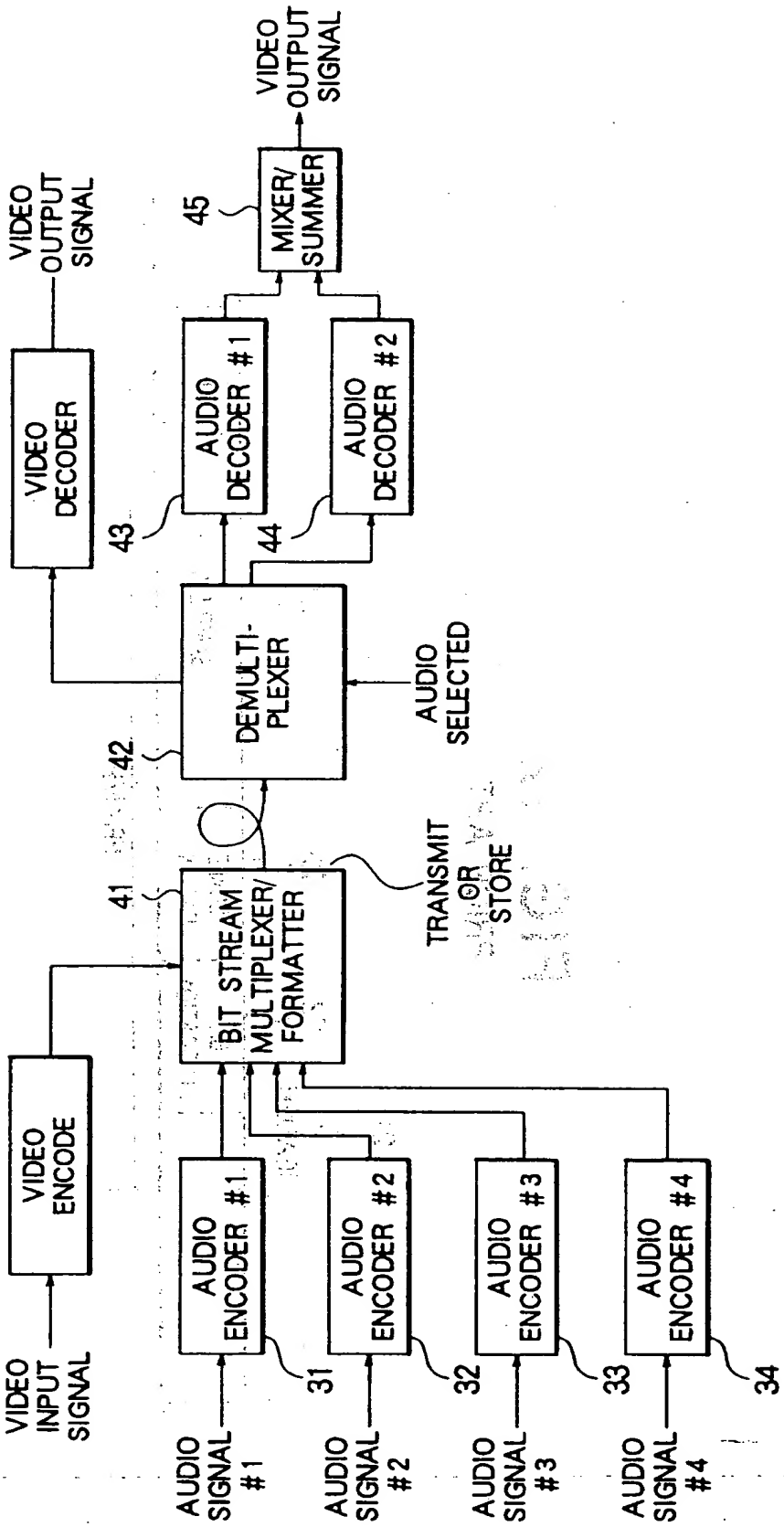


FIG. 13
PRIOR ART



(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 673 167 A3

# (12) EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
**31.07.1996 Bulletin 1996/31**

(43) Date of publication A2:
**20.09.1995 Bulletin 1995/38**

(21) Application number: **94118301.4**

(22) Date of filing: **21.11.1994**

(51) Int. Cl.[6]: **H04N 7/24**, H04N 7/52, H04N 5/60, G11B 20/10

(84) Designated Contracting States:
**DE FR GB NL**

(30) Priority: **16.03.1994 JP 46008/94**

(71) Applicant: **KABUSHIKI KAISHA TOSHIBA**
**Kawasaki-shi, Kanagawa-ken 210 (JP)**

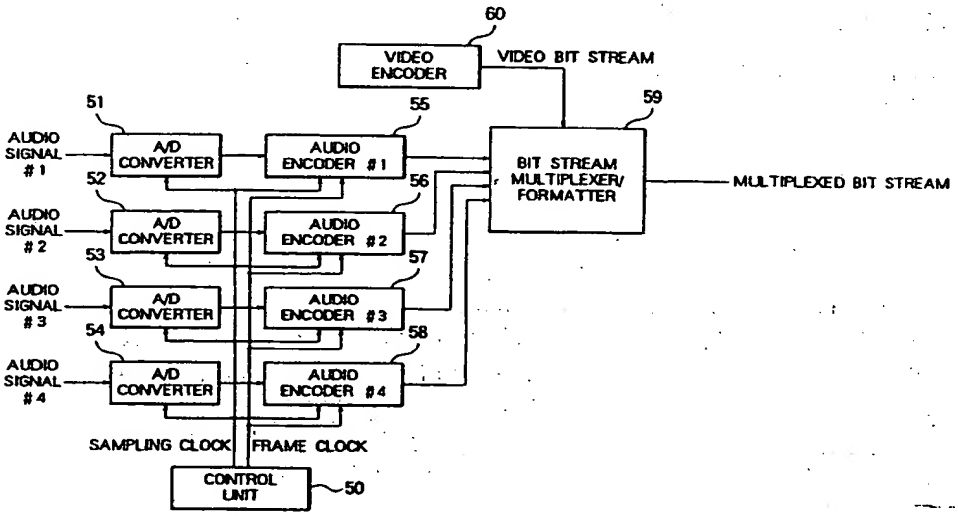
(72) Inventor: **Nakamura, Shin-ichi,**
**c/o Intellectual Prop. Div.**
**Minato-ku, Tokyo (JP)**

(74) Representative: **Henkel, Feiler, Hänzel & Partner**
**Möhlstrasse 37**
**81675 München (DE)**

## (54) Signal encoding and decoding system allowing adding of signals in a form of frequency sample sequence upon decoding

(57) In a signal compressing and encoding apparatus, a plurality of A/D converters sample input audio signals synchronously to produce time sample strings. An audio encoding circuit transforms time sample sequences in time/frequency by a frame interval to produce frequency sample sequences. The frequency sample sequences are compressed, encoded and formatted into a frame structure to generate audio bit streams. The audio bit streams are multiplexed by a formatter into a multiplexed bit stream. On the other hand, in a compressed signal decoding circuit, a quantization decoder decodes the frequency sample sequences for each bit stream separated and extracted selectively out of the multiplexed bit stream. A adding result the decoded frequency sample sequences to produce a summing result. The summing result is then transformed in frequency/time by a frequency/time mapping to produce the time sample sequences.

FIG. 1




EP 0 673 167 A3

European Patent Office

# EUROPEAN SEARCH REPORT

Application Number
EP 94 11 8301

## DOCUMENTS CONSIDERED TO BE RELEVANT

| Category | Citation of document with indication, where appropriate, of relevant passages | Relevant to claim | CLASSIFICATION OF THE APPLICATION (Int.Cl.6) |
|---|---|---|---|
| A | EP-A-0 521 487 (SONY CORP.)<br>* column 3, line 40 - column 4, line 16 *<br>* column 13, line 57 - column 14, line 54; figure 1 * | 1-12 | H04N7/24<br>H04N7/52<br>H04N5/60<br>G11B20/10 |
| A | DIGITAL VIDEO COMPRESSION ON PERSONAL COMPUTERS: ALGORITHMS AND TECHNOLOGIES, vol. 2187, 7 February 1994, SPIE, SAN JOSE, CAL, USA, pages 260-273, XP000571198<br>PAN: "An Overview of the MPEG/Audio Compression Algorithm"<br>* paragraph 3; figure 1 * | 1-12 | |
| A | SIGNAL PROCESSING: IMAGE COMMUNICATION, vol. 4, 1992, ELSEVIER, AMS, NL, pages 153-159, XP000273161<br>MAC INNIS: "The MPEG Systems Coding Specification"<br>* the whole document * | 1-12 | |
| A | US-A-5 130 815 (SILVERMAN ET AL.)<br>* abstract; claims; figure 1 * | 1-12 | TECHNICAL FIELDS SEARCHED (Int.Cl.6)<br>H04N<br>G11B |
| A | EP-A-0 381 807 (PIONEER ELECTRONIC CORP.)<br>* column 6, line 46 - column 7, line 16; claims; figure 4 * | 1-12 | |
| A | EP-A-0 372 155 (KARAMON ET AL.) | | |
| A | EP-A-0 245 904 (N.V. PHILIPS' GLOEILAMPENFABRIEKEN) | | |
| A | EP-A-0 577 329 (AT&T CO)<br>* abstract; figure 2 * | 1-12 | |

-/--

The present search report has been drawn up for all claims

| Place of search | Date of completion of the search | Examiner |
|---|---|---|
| THE HAGUE | 30 May 1996 | Foglia, P |

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

European Patent Office

# EUROPEAN SEARCH REPORT

Application Number
EP 94 11 8301

## DOCUMENTS CONSIDERED TO BE RELEVANT

| Category | Citation of document with indication, where appropriate, of relevant passages | Relevant to claim | CLASSIFICATION OF THE APPLICATION (Int.Cl.6) |
|---|---|---|---|
| A | PATENT ABSTRACTS OF JAPAN vol. 16, no. 303 (E-1228), 3 July 1992 & JP-A-04 082433 (NEC CORP. ET AL.), 16 March 1992, * abstract * --- | | |
| A | PATENT ABSTRACTS OF JAPAN vol. 14, no. 187 (E-917), 16 April 1990 & JP-A-02 035886 (FUJITSU LTD), 6 February 1990, * abstract * ----- | | |

TECHNICAL FIELDS SEARCHED (Int.Cl.6)

The present search report has been drawn up for all claims

| Place of search | Date of completion of the search | Examiner |
|---|---|---|
| THE HAGUE | 30 May 1996 | Foglia, P |

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)